Observing System Simulation Experiments: an overview

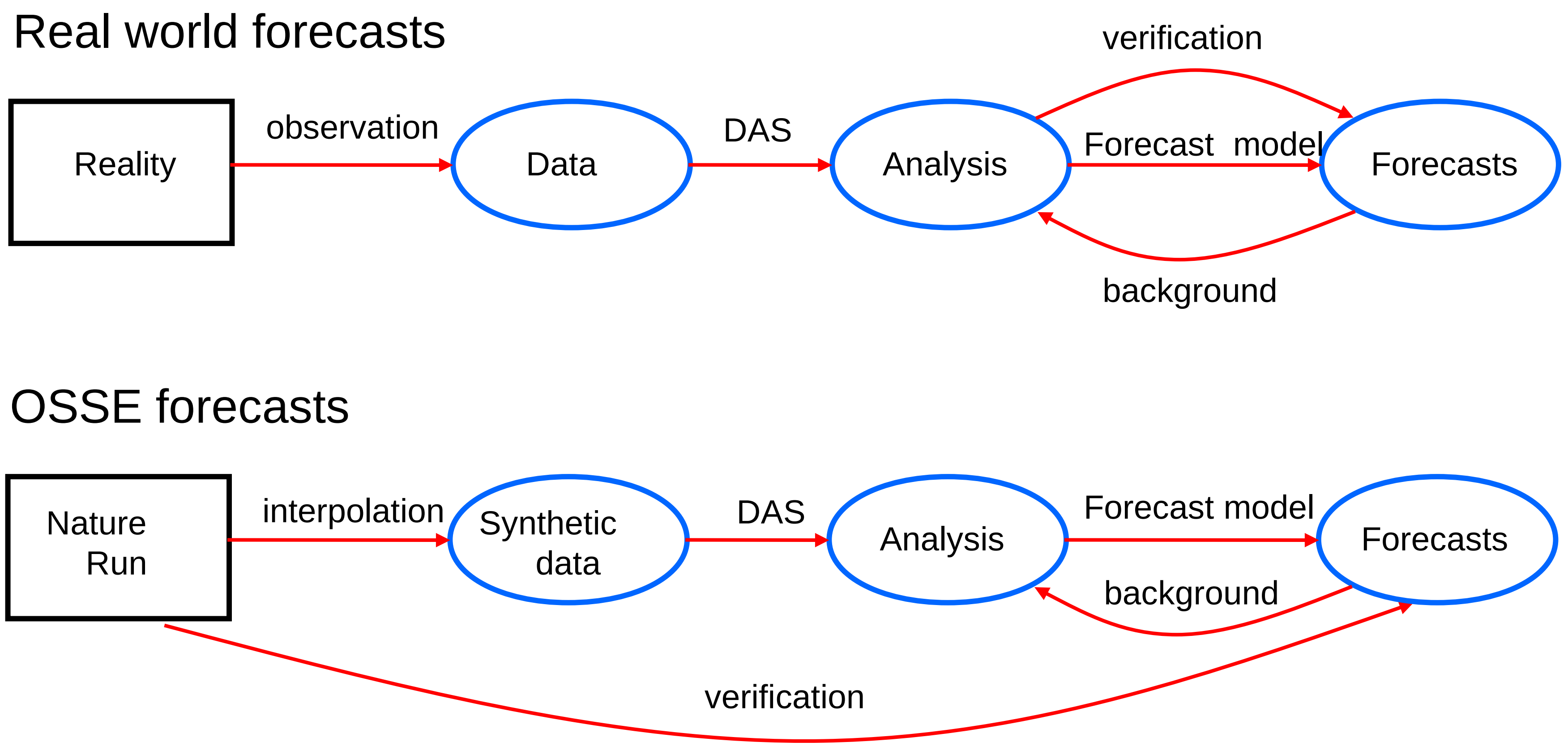
Nikki C. Privé
Ronald M. Errico

What is an OSSE?

An OSSE is a modeling experiment used to evaluate the impact of new observing systems on operational forecasts when actual observational data is not available.

- A long free model run is used as the "truth" - the Nature Run
- The Nature Run fields are used to back out "synthetic observations" from all current and new observing systems.
- Suitable errors are added to the synthetic observations
- The synthetic observations are assimilated into a different operational model
- Forecasts are made with the second model and compared with the Nature Run to quantify improvements due to the new observing system

OSSEs vs. the Real World



Why do an OSSE?

1. You want to find out if a new observing system will add value to NWP analyses and forecasts

2. You want to make design decisions for a new observing system

3. You want to investigate the behavior of data assimilation systems in an environment where the truth is known

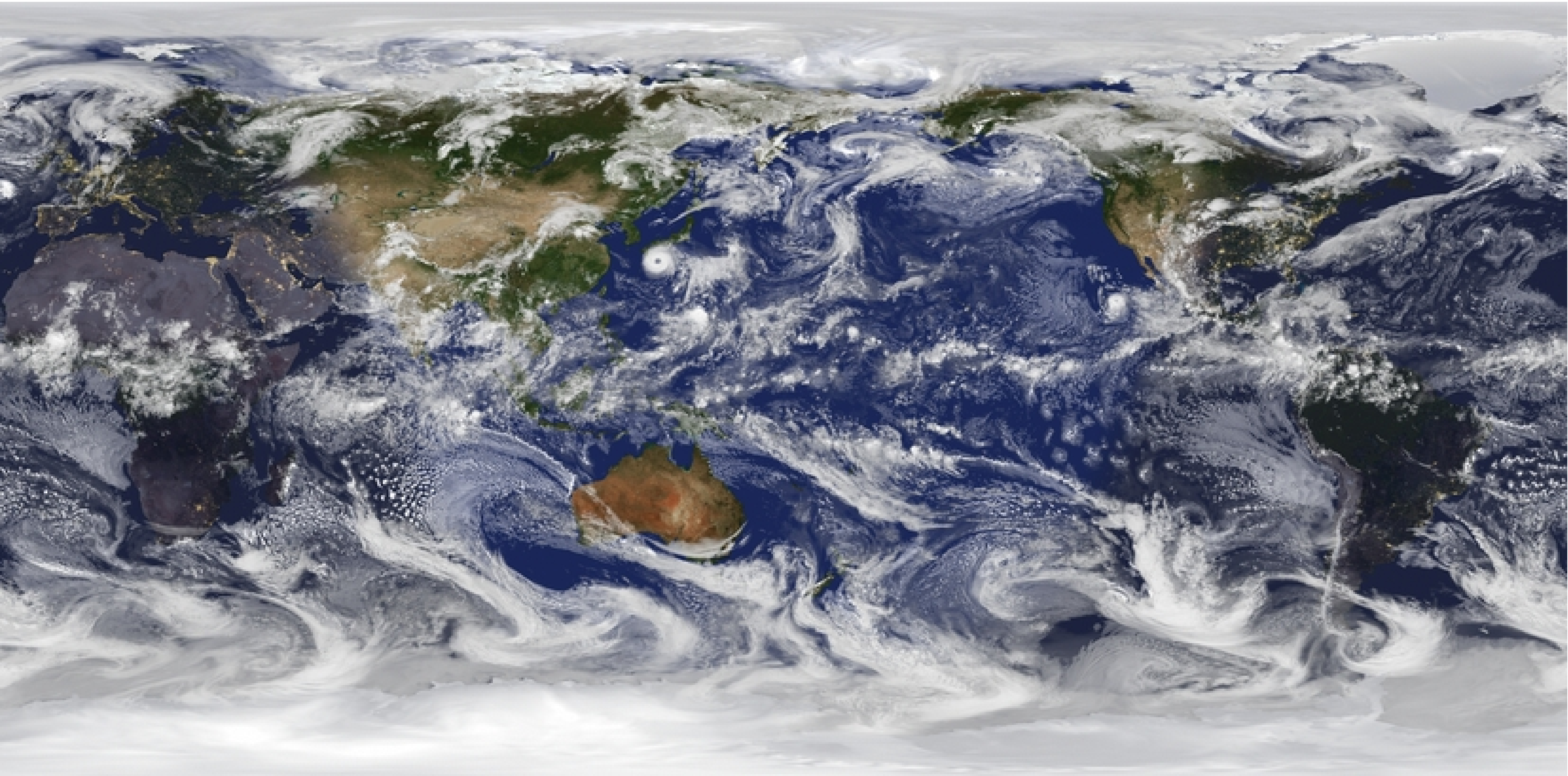
When not to run an OSSE

- When you can't model the phenomena you are interested in
- When you can't simulate your new observations
- When you can't assimilate your new observations

Nature Runs

- Nature Runs act as the 'truth' in the OSSE, replacing the real atmosphere.
- Usually, a long free (non-cycling) forecast from the best available model is used as the NR
  - Model forecast has continuity of fields in time
  - Sometimes an analysis or reanalysis sequence is used, but the sequence of states of truth can never be replicated by a model
  - Always a push for bigger, higher resolution NR

G5 Nature Run



2 year, 7 km/72L, 30 minute resolution

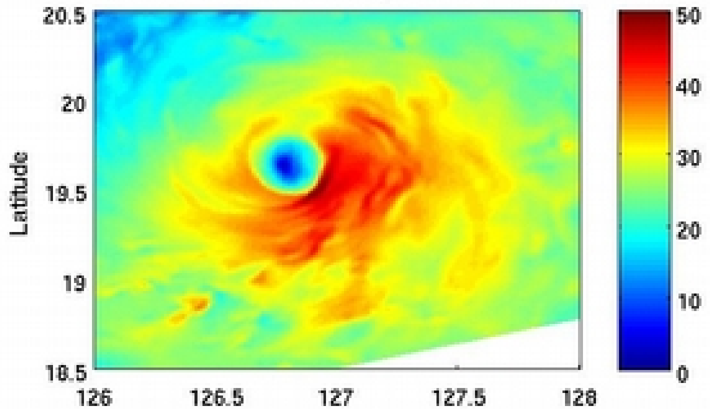
15 aerosols, ozone, CO, CO_2

Image courtesy: William Putman

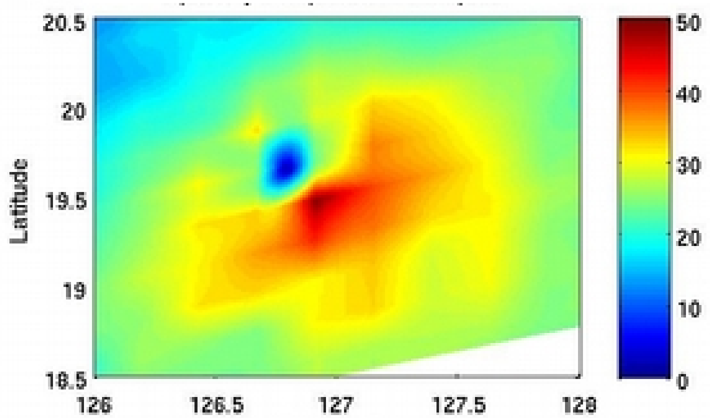
Nature Run Requirements

- Must be able to realistically model phenomena of interest
  - Dynamics and physics should be realistic
  - Must produce fields needed for “observations”
  - Should be verified against real world
- Ideally is ‘better’ than the operational model to be used for experiments
- Preferably a different model base is used for the NR and the experimental forecast model to reduce incestuousness

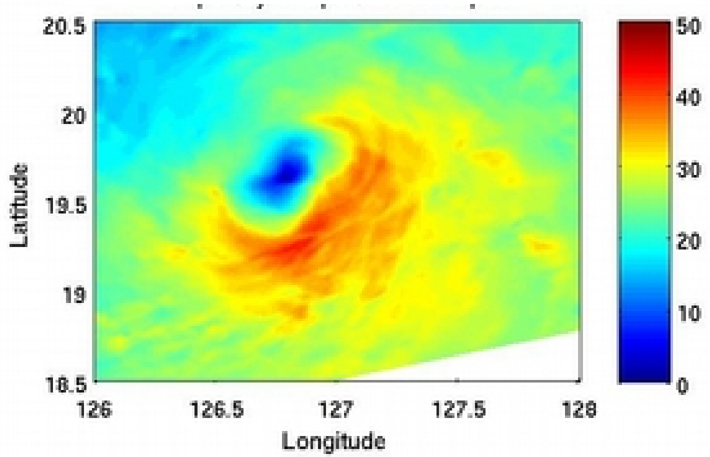
True Windspeed, 1.5 km, 10 min



Spatial Interpolation, 27 km



Temporal Interpolation, 180 min



Lessons Learned: Nature Run

- Higher spatial resolution is not always sufficient
  - Temporal output needs to keep up with spatial output
  - Large datasets are difficult to store and handle

Privé, N. C., and R. M. Errico, 2016. Temporal and spatial interpolation errors of high-resolution modeled atmospheric fields. *J. Atmos. Ocean. Tech.*, in press.

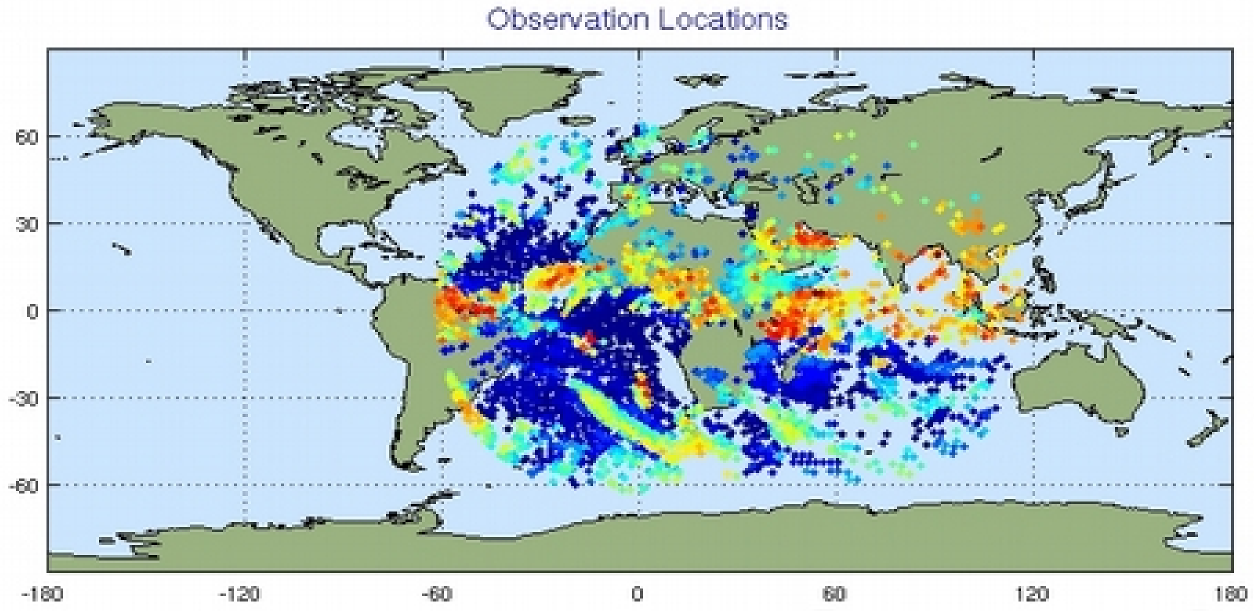
Nature Run Validation

- Evaluate if NR is sufficiently realistic to yield meaningful results
- In addition to the phenomena of interest, the NR needs to realistically replicate fields needed to generate synthetic observations
- Can't validate everything; corollary – don't expect a NR to come pre-validated for your needs

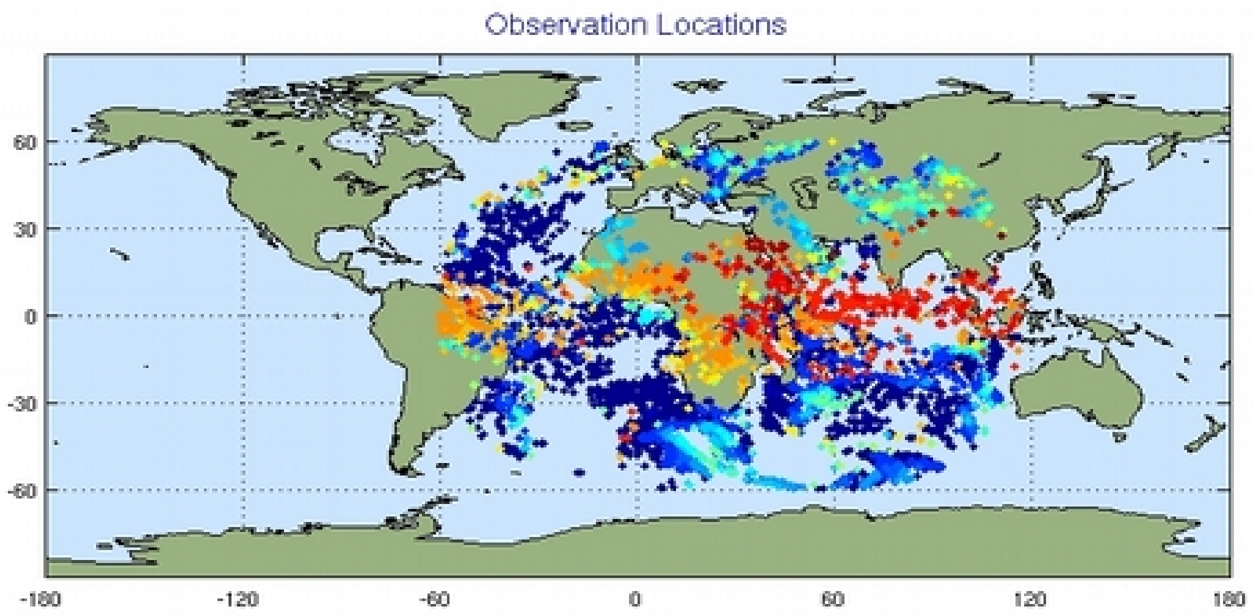
Synthetic Observations

Example of METEOSAT AMV observations at 00 UTC 10 July

Real



Simulated

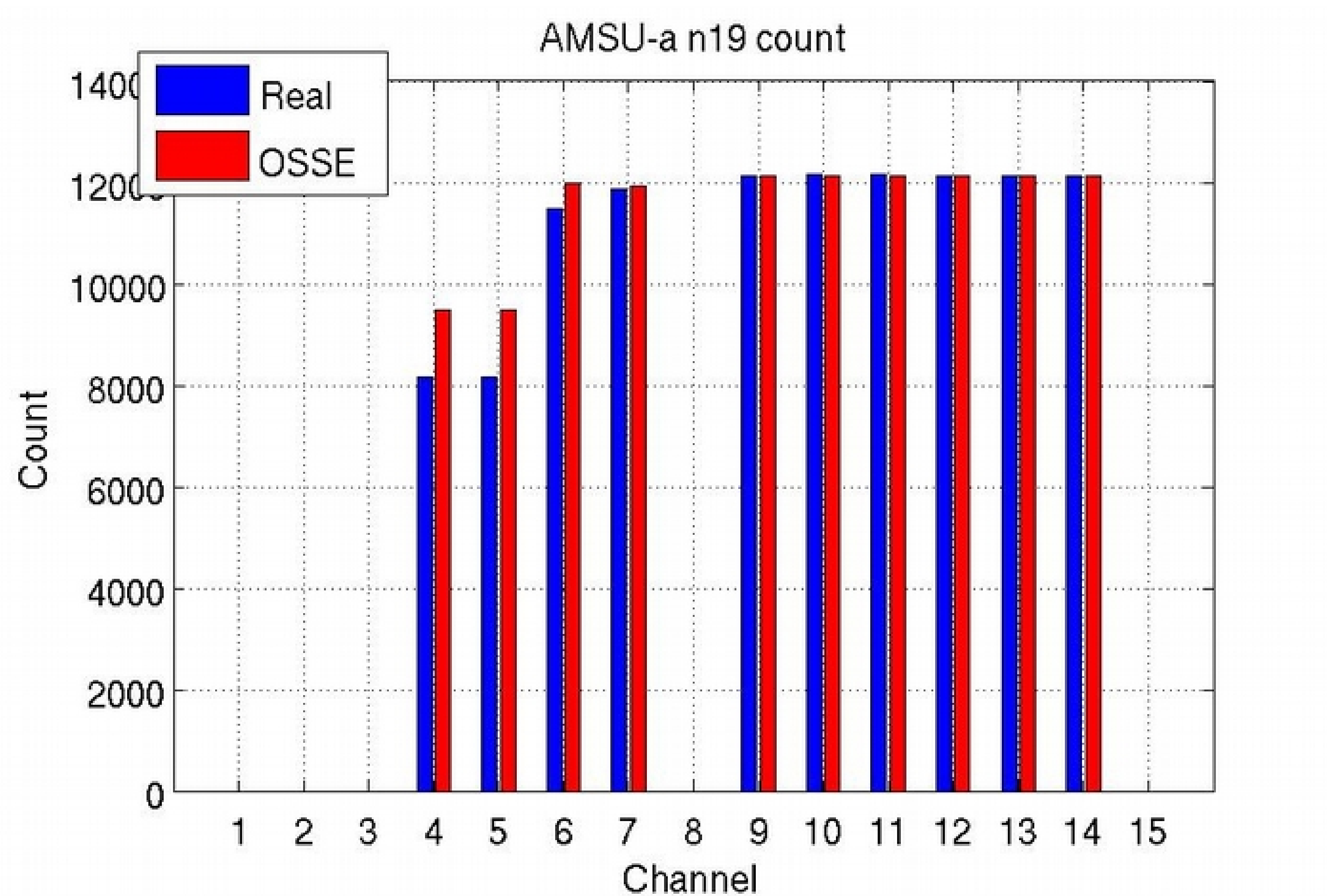


Observation Errors

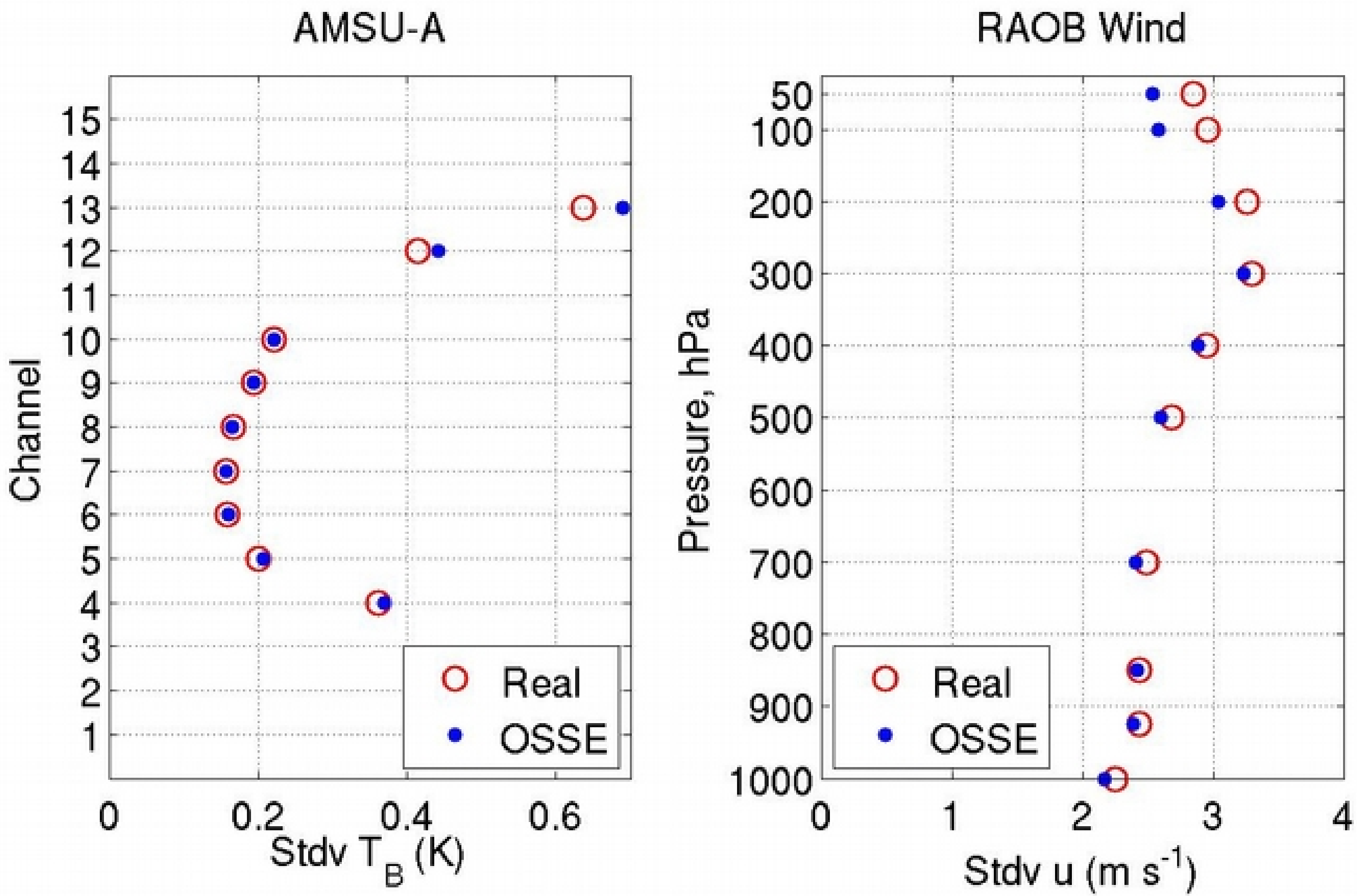
- Synthetic observations contain some intrinsic interpolation/operator errors, but less than real observations (usually)
- Synthetic errors are created and added to the synthetic observations to compensate
- Error is complex and poorly understood
  - Error magnitude
  - Biases
  - Correlated errors

Calibration

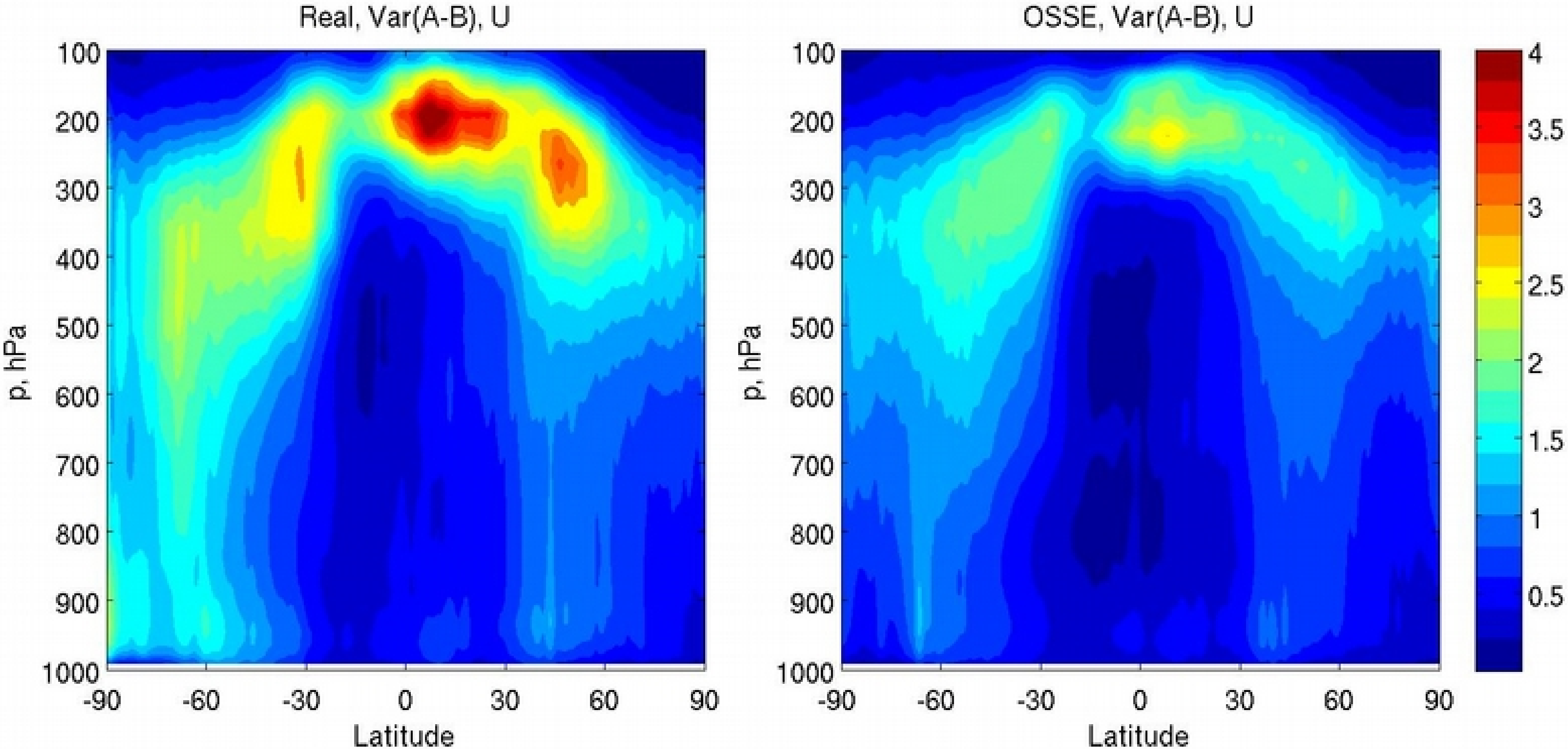
- Adjust synthetic observations and their errors to increase realism of the OSSE in a statistical sense
  - Compare OSSE statistics to statistics using real data in the same DAS/forecast system
- Need to decide what statistical metrics to use for the calibration, depending on your needs
- Calibrating new observation types?
  - Find an analogous data type if possible



- Observation count is easy to calibrate



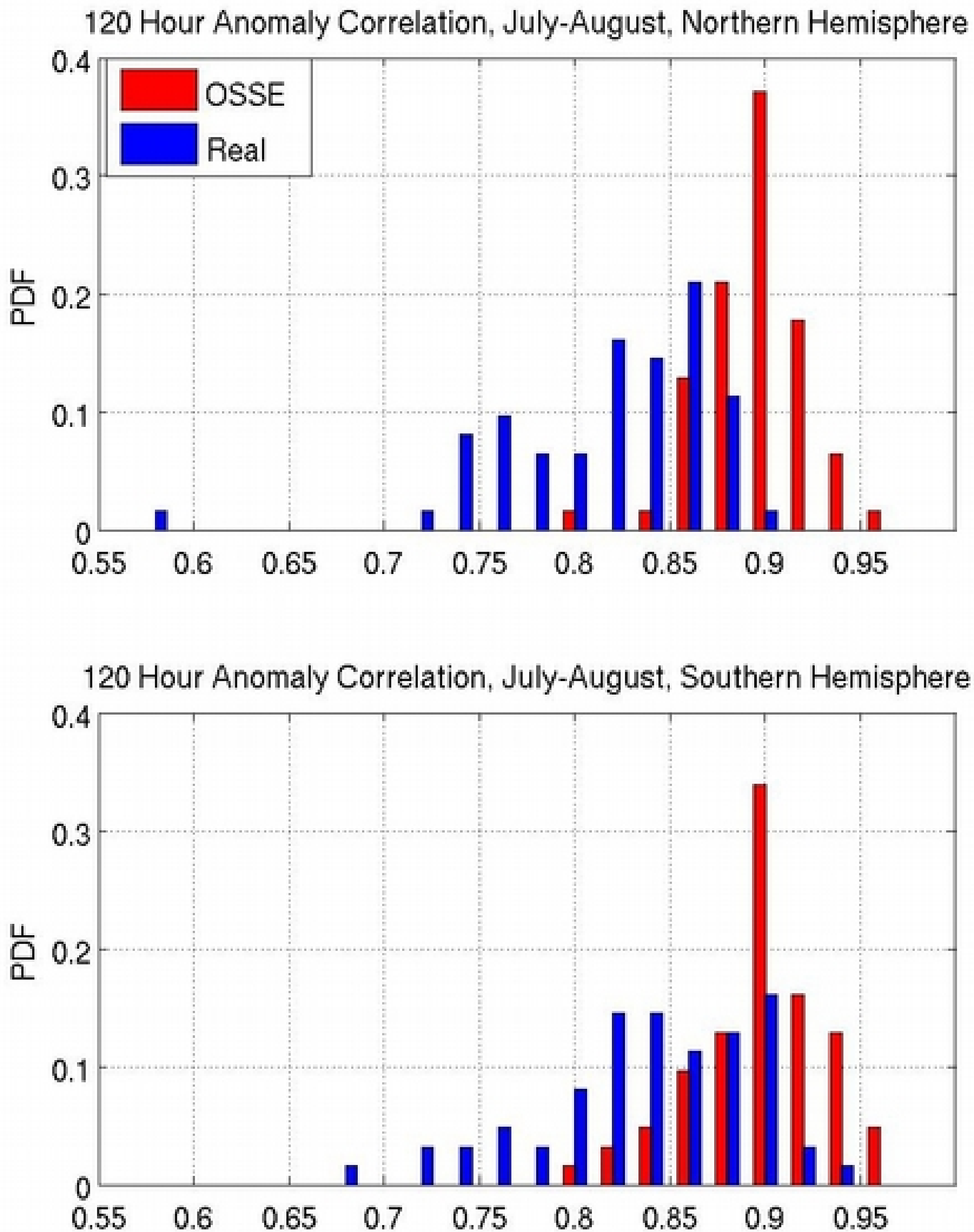
O-F is fairly easy to calibrate because you can manipulate O directly.



A-B (analysis increment) is a little harder to calibrate, as A and B are not directly controlled

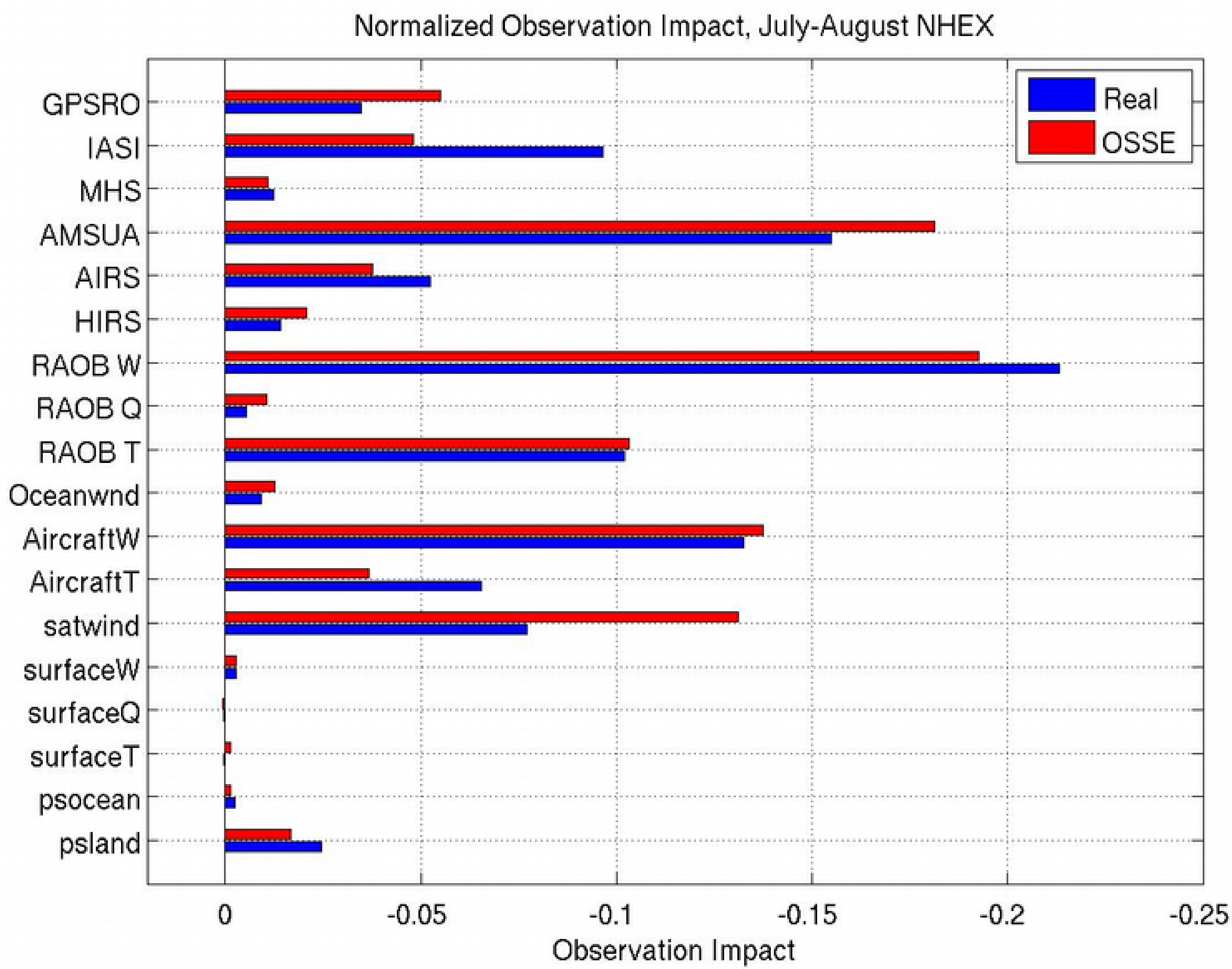
Zonal mean monthly temporal variances of (A-B) for zonal wind, G5NR

Model error strongly influences forecast skill in the longer term forecast, so calibration is not possible (unless you want to mess with your model).



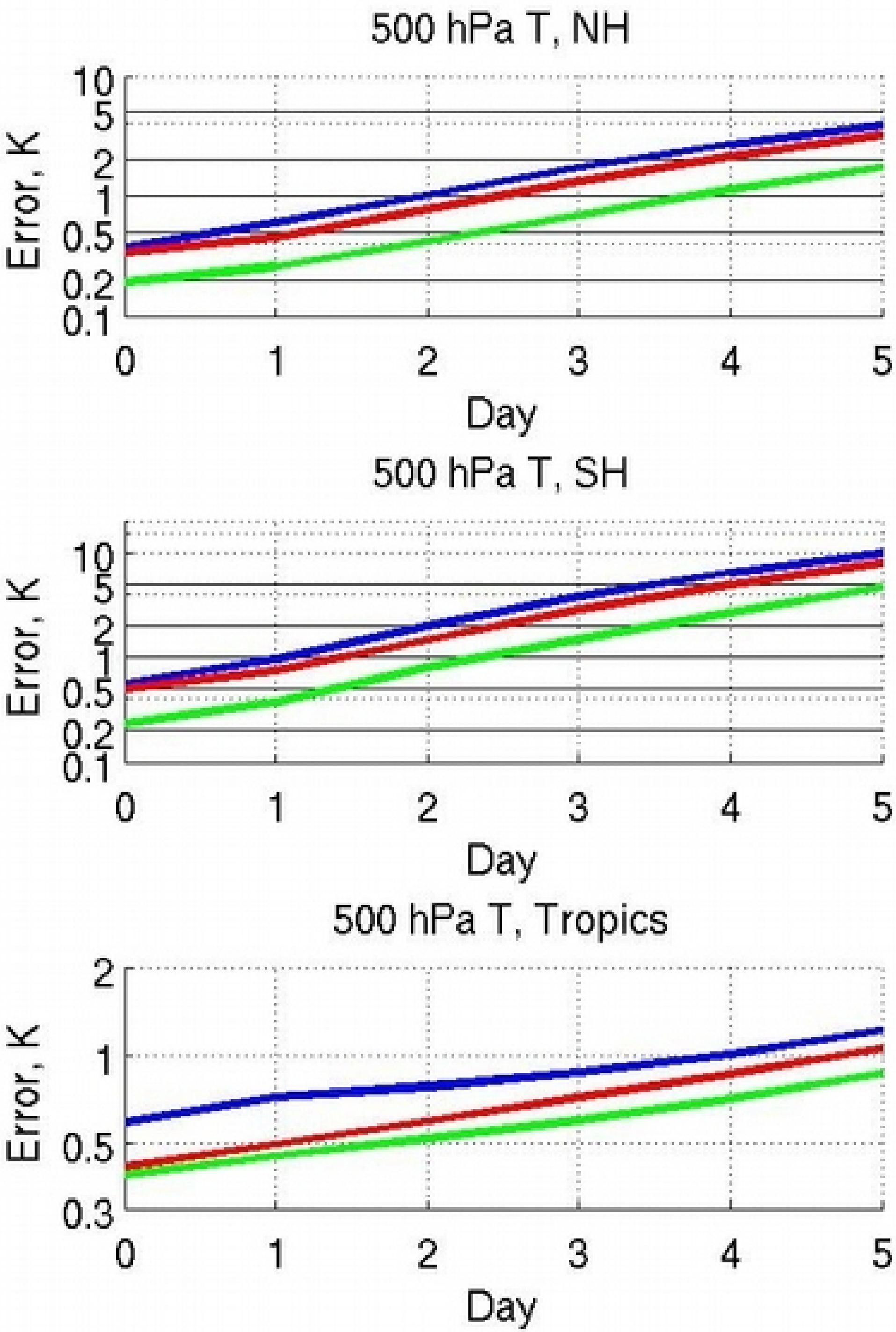
Red: OSSE
Blue: Real

500 hPa anomaly correlations of geopotential height, G5NR



Normalized Observation Impact, July-August NHEX
Real
OSSE
GPSRO
IASI
MHS
AMSUA
AIRS
HIRS
RAOB W
RAOB Q
RAOB T
Oceanwnd
AircraftW
AircraftT
satwind
surfaceW
surfaceQ
surfaceT
psocean
psland
0
-0.05
-0.1
-0.15
-0.2
-0.25
Observation Impact

500 hPa error variance for temperature as a function of forecast time



Red: Fraternal-Twin OSSE (G5NR)
Blue: Non-Twin OSSE (ECMWF T511 NR)
Green: Identical Twin OSSE (GEOS-5 self NR)

Fraternal twin OSSE has forecast error closer to non-twin OSSE in the extratropics.

A range of behaviors are observed in the tropics.

Choosing Metrics

- Long cycling periods necessary to get statistically significant results for most new observations
- Anomaly correlation is a difficult metric to show appreciable impacts
- What fields do you expect the instrument to improve?
- Largest impacts found at analysis time or short-term forecasts

Criticisms and Pitfalls of OSSEs

- Results only apply within the OSSE system – no concrete connection to the real world
- Even the best OSSEs are far from perfect: incestuousness, difficulty in generating observations and errors, deficiencies of the Nature Run
- By the time the new instrument is deployed, both the global observing network and the forecast models/DAS will be different
  - Examples of sloppy or unsuccessful OSSEs
  - Very reduced baseline of assimilated observational data (ex. no radiance data)
  - Other artificial degradation of analysis state
  - No validation or calibration of OSSE framework

Takeaways

- OSSEs can provide useful information about new observational types and the workings of data assimilation systems
- Careful consideration of research goals should guide each step of the OSSE process
- OSSEs are hard, good OSSEs are harder